Real-time detection of the start and subsequent epidemic states of HIV outbreaks among people who inject drugs: insights from four European countries

Valia Baralou[1], Argiro Karakosta[1], Christos Thomadakis[1], Nikos Demiris[2], Nikos Pantazis[1], Olga Anagnostou[3], Christos Danopoulos[3], Dimitris Katsiris[4], and Giota Touloumi[1]

[1]Department of Hygiene, Epidemiology & Medical Statistics, Medical School, National & Kapodistrian University of Athens, Athens, Greece
[2]Department of Statistics, Athens University of Economics and Business, Athens, Greece
[3]Greek Organisation Against Drugs (OKANA), Athens, Greece
[4]InDigital S.A., Athens, Greece

May 09, 2023

Injecting drug use and HIV in Europe

- HIV outbreaks among people who inject drugs (PWID) have occurred in many European countries over the past decade with serious impact on public health (Des Jarlais et al., 2020).

Injecting drug use and HIV in Europe

- HIV outbreaks among people who inject drugs (PWID) have occurred in many European countries over the past decade with serious impact on public health (Des Jarlais et al., 2020).
- No automated real-time outbreak detection system has been developed for monitoring HIV among PWID yet.

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).
  - Classic methods

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).
  - Classic methods
    - Control charts
    - Regression-based methods

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).
  - Classic methods
    - Control charts
    - Regression-based methods
  - Novel approaches

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).
  - Classic methods
    - Control charts
    - Regression-based methods
  - Novel approaches
    - Hidden Markov Model (HMM)
    - Bayesian methods

Methods for real-time outbreak detection

- A broad range of prospective outbreak detection methods have been proposed in literature (Yuan et al., 2019).
  - Classic methods
    - Control charts
    - Regression-based methods
  - Novel approaches
    - Hidden Markov Model (HMM)
    - Bayesian methods
- Although novel methods seem promising, most comparative studies focused on the classic ones and were based on simulations.

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).
- Unlike other epidemics widely investigated, HIV outbreaks among PWID differ in their:

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).
- Unlike other epidemics widely investigated, HIV outbreaks among PWID differ in their:
  - size

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).
- Unlike other epidemics widely investigated, HIV outbreaks among PWID differ in their:
  - size
  - duration and

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).
- Unlike other epidemics widely investigated, HIV outbreaks among PWID differ in their:
  - size
  - duration and
  - post-epidemic level.

Gap in literature

- Due to limited data availability, the investigation of real outbreaks is underexplored but may provide useful insights that cannot be revealed by simulated data.
- Only a few studies have intended to identify epidemic states after the outbreak onset while all but one addressed the problem retrospectively rather than prospectively (Lytras et al., 2019; Jobart et al., 2021).
- Unlike other epidemics widely investigated, HIV outbreaks among PWID differ in their:
    - ◇ size
    - ◇ duration and
    - ◇ post-epidemic level.
- Research questions on how anomaly detection methods behave when applied to such epidemic curves remain still open.

Aim

In this study, we compared the performance of different methods for detecting in real time the:

Aim

In this study, we compared the performance of different methods for detecting in real time the:

- growth

Aim

In this study, we compared the performance of different methods for detecting in real time the:

- growth
- decline and

Aim

In this study, we compared the performance of different methods for detecting in real time the:

- growth
- decline and
- post-epidemic state

of HIV outbreaks among PWID in four European countries.

Methods: Overview (1)

We used the following methods to identify each epidemic state:

Methods: Overview (1)

We used the following methods to identify each epidemic state:

- a two-state HMM

Methods: Overview (1)

We used the following methods to identify each epidemic state:

- a two-state HMM
- a new method based on prediction interval (PI)

Methods: Overview (1)

We used the following methods to identify each epidemic state:

- a two-state HMM
- a new method based on prediction interval (PI)
- a novel approach HMM-PI that combines the former two

Methods: Overview (1)

We used the following methods to identify each epidemic state:

- a two-state HMM
- a new method based on prediction interval (PI)
- a novel approach HMM-PI that combines the former two
- control charts for outbreak detection
  - ◇ Exponentially Weighted Moving Average (EWMA) and
  - ◇ Cumulative Sum (CUSUM) chart for Poisson variate.

Methods: Overview (2)

Detection of epidemic states was sequentially conducted

Pre-epidemic state → **Growth** → **Decline** → **Post-epidemic state**

Methods: Overview (2)

Detection of epidemic states was sequentially conducted

Pre-epidemic state → **Growth** → **Decline** → **Post-epidemic state**

Baseline (i.e., counts without outbreaks): historical pre-epidemic cases extending three years in total

Methods: Overview (2)

Detection of epidemic states was sequentially conducted

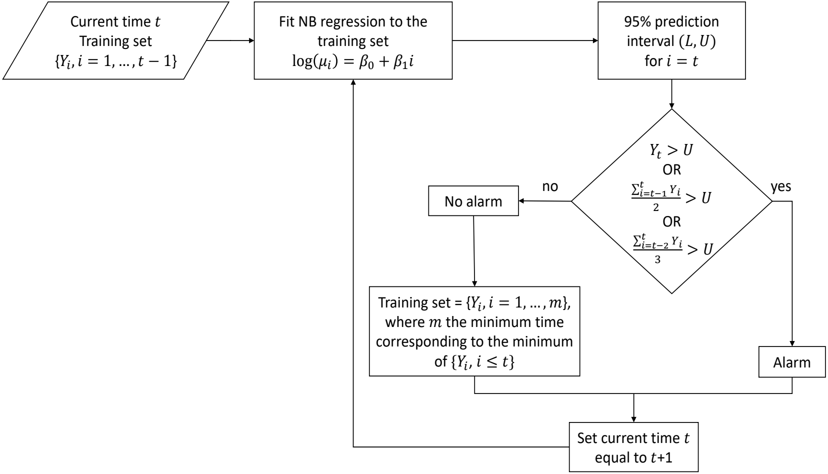
Pre-epidemic state → **Growth** → **Decline** → **Post-epidemic state**

Baseline (i.e., counts without outbreaks): historical pre-epidemic cases extending three years in total

Constraint: ≥ 12 observations in training set i.e., once the epidemic state changed, it remained unchanged for the following 11 weeks

Growth phase: Method based on prediction interval (PI)

Schematic outline of the PI method for outbreak detection



Growth phase: Hidden Markov Model

- We used a two-state HMM to estimate the transition probability from pre-epidemic $(Z_t = 1)$ to growth $(Z_t = 2)$ state at each time point; with the following mathematical form:

$$\begin{pmatrix} 1-p_{12} & p_{12} \\ 0 & 1 \end{pmatrix} \text{ and } Y_t \sim NB(\mu_t, r_t), \text{ where}$$

$$\mu_{t|Z_{t=1}} = \mu_{t-1} \text{ and } \mu_{t|Z_{t=2}} = \mu_{t-1}e^{\beta_1},\ \beta_1 > 0$$

$$r_{t|Z_{t=1}} = r_1 \text{ and } r_{t|Z_{t=2}} = r_2 \text{ such that } r_1 < r_2$$

Growth phase: Hidden Markov Model

- We used a two-state HMM to estimate the transition probability from pre-epidemic ($Z_t = 1$) to growth ($Z_t = 2$) state at each time point; with the following mathematical form:

$$\begin{pmatrix} 1-p_{12} & p_{12} \\ 0 & 1 \end{pmatrix} \text{ and } Y_t \sim NB(\mu_t, r_t), \text{ where}$$

$$\mu_{t|Z_{t=1}} = \mu_{t-1} \text{ and } \mu_{t|Z_{t=2}} = \mu_{t-1} e^{\beta_1},\ \beta_1 > 0$$

$$r_{t|Z_{t=1}} = r_1 \text{ and } r_{t|Z_{t=2}} = r_2 \text{ such that } r_1 < r_2$$

- HMM was fitted in a Bayesian framework using Markov Chain Monte Carlo.

Growth phase: Hidden Markov Model

- We used a two-state HMM to estimate the transition probability from pre-epidemic $(Z_t = 1)$ to growth $(Z_t = 2)$ state at each time point; with the following mathematical form:

$$\begin{pmatrix} 1 - p_{12} & p_{12} \\ 0 & 1 \end{pmatrix} \text{ and } Y_t \sim NB(\mu_t, r_t), \text{ where}$$

$$\mu_{t|Z_t=1} = \mu_{t-1} \text{ and } \mu_{t|Z_t=2} = \mu_{t-1} e^{\beta_1},\ \beta_1 > 0$$

$$r_{t|Z_t=1} = r_1 \text{ and } r_{t|Z_t=2} = r_2 \text{ such that } r_1 < r_2$$

- HMM was fitted in a Bayesian framework using Markov Chain Monte Carlo.
- An alert was raised when the posterior probability of transitioning from the pre-epidemic state to growth exceeded 50%.

Growth phase: HMM-PI approach

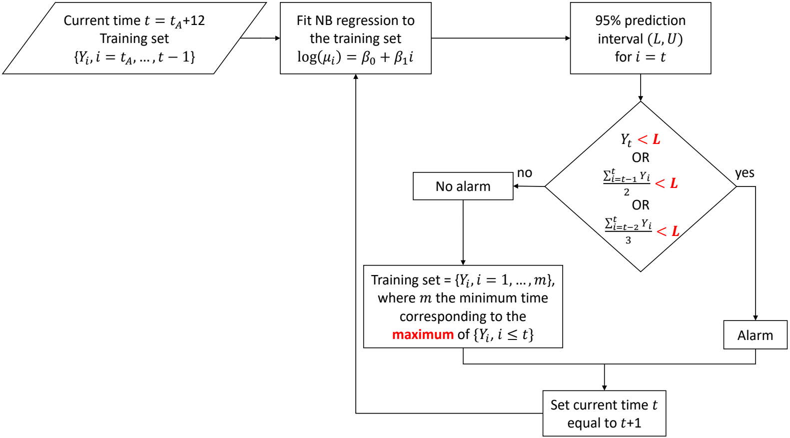
We proposed a new method, called HMM-PI, that combines HMM and PI by assigning weights to HMM's transition probabilities at each week based on the PI upper limit as follows:

$$\begin{pmatrix} w_t(1-p_{12}) & 1-w_t(1-p_{12}) \\ 0 & 1 \end{pmatrix}, \text{ where } w_t = \begin{cases} \frac{Y_t - U}{Y_t} & , Y_t > U \\ 1 & , Y_t \leq U \end{cases}$$

The weighting scheme was applied for a fixed time horizon (i.e, 12 weeks) after four consecutive alarm signals triggered by PI.

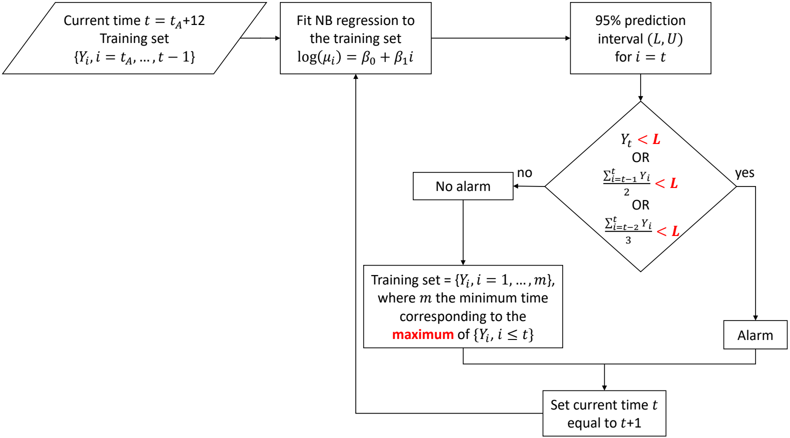
Decline and post-epidemic state: PI method

Schematic outline of the PI method for **decline** detection.



Decline and post-epidemic state: PI method

Schematic outline of the PI method for **decline** detection.



To identify the **post-epidemic** state, we used the PI method for outbreak detection.

Decline and post-epidemic state: HMM and HMM-PI

- Similarly with growth, we used a two-state HMM with the following changes in its mathematical form to estimate the transition probability:

Decline and post-epidemic state: HMM and HMM-PI

- Similarly with growth, we used a two-state HMM with the following changes in its mathematical form to estimate the transition probability:
  - **from growth** $(Z_t = 2)$ **to decline** $(Z_t = 3)$ phase

$$\mu_{t|Z_t=2} = \mu_{t-1}e^{\beta_1},\ \beta_1 > 0 \text{ and } \mu_{t|Z_t=3} = \mu_{t-1}e^{\beta_2},\ \beta_2 < 0$$

$$r_{t|Z_t=2} = r_{t|Z_t=3} = r$$

Decline and post-epidemic state: HMM and HMM-PI

- Similarly with growth, we used a two-state HMM with the following changes in its mathematical form to estimate the transition probability:
  - **from growth** $(Z_t = 2)$ **to decline** $(Z_t = 3)$ phase

$$\mu_{t|Z_t=2} = \mu_{t-1}e^{\beta_1},\ \beta_1 > 0 \text{ and } \mu_{t|Z_t=3} = \mu_{t-1}e^{\beta_2},\ \beta_2 < 0$$

$$r_{t|Z_t=2} = r_{t|Z_t=3} = r$$

  - **from decline** $(Z_t = 3)$ **to post-epidemic** $(Z_t = 4)$ state

$$\mu_{t|Z_t=3} = \mu_{t-1}e^{\beta_2},\ \beta_2 < 0 \text{ and } \mu_{t|Z_t=4} = \mu_{t-1}$$

$$r_{t|Z_t=3} = r_3 \text{ and } r_{t|Z_t=4} = r_4 \text{ such that } r_3 < r_4$$

Decline and post-epidemic state: HMM and HMM-PI

- Similarly with growth, we used a two-state HMM with the following changes in its mathematical form to estimate the transition probability:
  - **from growth** $(Z_t = 2)$ **to decline** $(Z_t = 3)$ phase

$$\mu_{t|Z_t=2} = \mu_{t-1}e^{\beta_1},\ \beta_1 > 0 \text{ and } \mu_{t|Z_t=3} = \mu_{t-1}e^{\beta_2},\ \beta_2 < 0$$

$$r_{t|Z_t=2} = r_{t|Z_t=3} = r$$

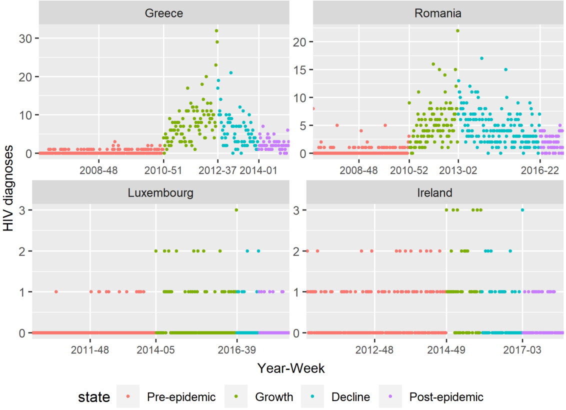
  - **from decline** $(Z_t = 3)$ **to post-epidemic** $(Z_t = 4)$ state

$$\mu_{t|Z_t=3} = \mu_{t-1}e^{\beta_2},\ \beta_2 < 0 \text{ and } \mu_{t|Z_t=4} = \mu_{t-1}$$

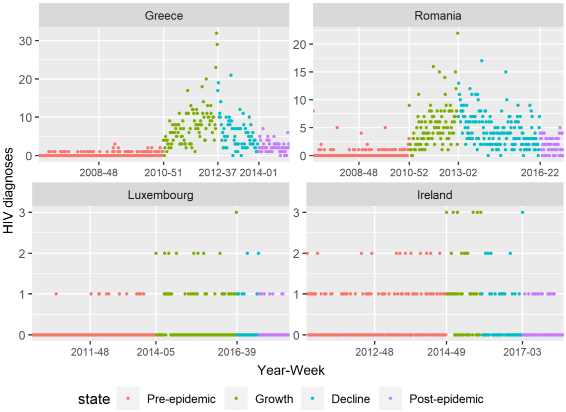
$$r_{t|Z_t=3} = r_3 \text{ and } r_{t|Z_t=4} = r_4 \text{ such that } r_3 < r_4$$

- HMM-PI combined HMM and PI integrating the appropriate transformations for the decline and post-epidemic state.

Data

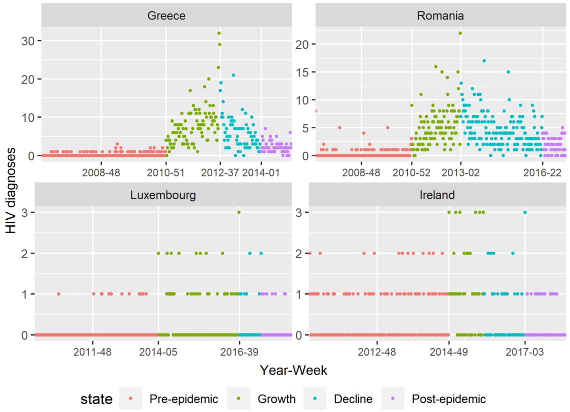

Greece
Romania
Luxembourg
Ireland
HIV diagnoses
Year-Week
2008-48
2010-51
2012-37
2014-01
2008-48
2010-52
2013-02
2016-22
2011-48
2014-05
2016-39
2012-48
2014-49
2017-03
state
Pre-epidemic
Growth
Decline
Post-epidemic

Data



- The exact duration of each state was determined by visual inspection and relevant literature.

Data



- The exact duration of each state was determined by visual inspection and relevant literature.
- Data were provided by The European Surveillance System (TESSy) of European Centre for Disease Prevention and Control (ECDC).

Performance metrics

- Methods were applied in real time as follows:

Pre-epidemic state → Growth → Decline → Post-epidemic state

Performance metrics

- Methods were applied in real time as follows:

Pre-epidemic state → Growth → Decline → Post-epidemic state

- Performance was assessed with three popular metrics:

Performance metrics

- Methods were applied in real time as follows:

Pre-epidemic state → Growth → Decline → Post-epidemic state

- Performance was assessed with three popular metrics:
  - ◇ sensitivity: true positive rate

$$\frac{Number\ of\ weeks\ in\ state\ z\ and\ associated\ with\ an\ alarm}{Number\ of\ weeks\ that\ are\ truly\ in\ state\ z}$$

Performance metrics

- Methods were applied in real time as follows:

Pre-epidemic state → Growth → Decline → Post-epidemic state

- Performance was assessed with three popular metrics:
  - ◇ sensitivity: true positive rate

$$\frac{Number\ of\ weeks\ in\ state\ z\ and\ associated\ with\ an\ alarm}{Number\ of\ weeks\ that\ are\ truly\ in\ state\ z}$$

  - ◇ specificity: true negative rate

$$\frac{Number\ of\ weeks\ not\ in\ state\ z\ and\ without\ an\ alarm}{Number\ of\ weeks\ that\ are\ not\ in\ state\ z}$$

Performance metrics

- Methods were applied in real time as follows:

Pre-epidemic state → Growth → Decline → Post-epidemic state

- Performance was assessed with three popular metrics:
  - sensitivity: true positive rate

$$\frac{Number\ of\ weeks\ in\ state\ z\ and\ associated\ with\ an\ alarm}{Number\ of\ weeks\ that\ are\ truly\ in\ state\ z}$$

  - specificity: true negative rate

$$\frac{Number\ of\ weeks\ not\ in\ state\ z\ and\ without\ an\ alarm}{Number\ of\ weeks\ that\ are\ not\ in\ state\ z}$$

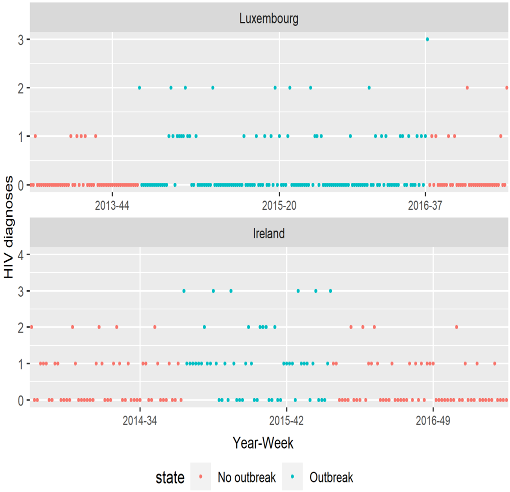
  - timeliness: difference between the start of each state and the first alarm signal after its onset.

Results: Detection of small outbreaks

Key findings

- Extremely small outbreaks that returned to pre-epidemic levels
- Only control charts identified the growth and post-epidemic state.

Country	Method	Sensitivity	Specificity
Luxembourg	CUSUM	94%	<50%
	EWMA	55%	69%
Ireland	CUSUM	57%	98%
	EWMA	88%	98%

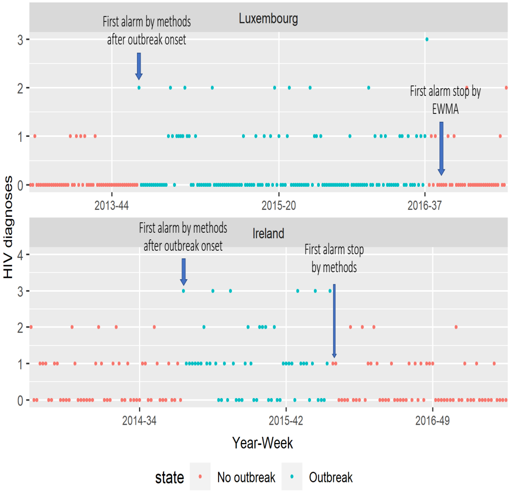


Results: Detection of small outbreaks

Key findings

- Extremely small outbreaks that returned to pre-epidemic levels
- Only control charts identified the growth and post-epidemic state.

Country	Method	Sensitivity	Specificity
Luxembourg	CUSUM	94%	<50%
	EWMA	55%	69%
Ireland	CUSUM	57%	98%
	EWMA	88%	98%

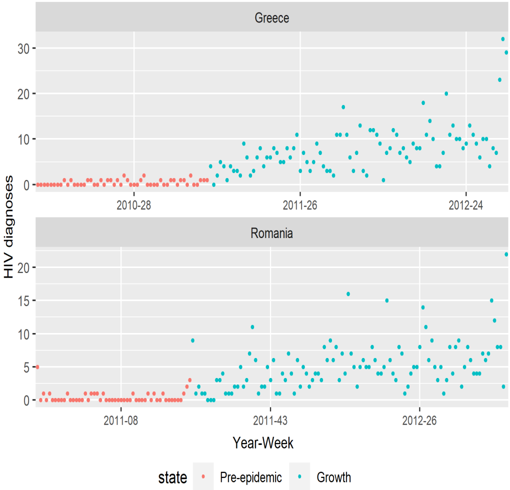


Results: Epidemic growth

Key findings

- Almost no delay in alarm
- PI performed at least similarly with control charts
- HMM-PI: excellent performance albeit with moderate specificity in Romania

Country	Method	Sensitivity	Specificity
Greece	CUSUM	100%	92%
	EWMA	100%	90%
	HMM	76%	100%
	PI	94%	100%
	HMM-PI	100%	100%
Romania	CUSUM	100%	85%
	EWMA	100%	90%
	HMM	59%	93%
	PI	87%	96%
	HMM-PI	100%	$\approx 50\%$

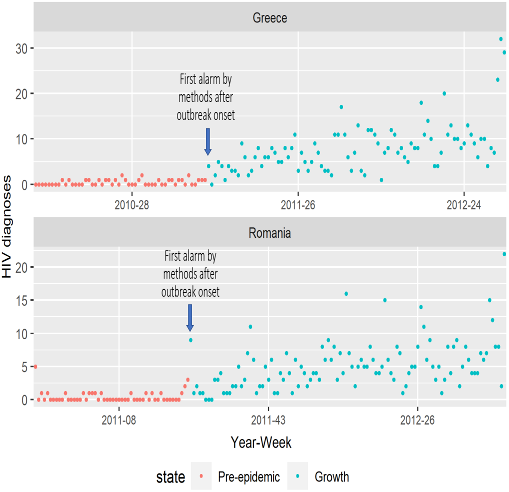


Results: Epidemic growth

Key findings

- Almost no delay in alarm
- PI performed at least similarly with control charts
- HMM-PI: excellent performance albeit with moderate specificity in Romania

Country	Method	Sensitivity	Specificity
Greece	CUSUM	100%	92%
	EWMA	100%	90%
	HMM	76%	100%
	PI	94%	100%
	HMM-PI	100%	100%
Romania	CUSUM	100%	85%
	EWMA	100%	90%
	HMM	59%	93%
	PI	87%	96%
	HMM-PI	100%	$\approx 50\%$



Results: Decline phase

- Post-epidemic counts in Romania and Greece were stabilised at a level higher than the pre-epidemic $\Rightarrow$ control charts failed.

Results: Decline phase

- Post-epidemic counts in Romania and Greece were stabilised at a level higher than the pre-epidemic ⇒ control charts failed.
- The remaining methods were applied for **decline** detection.

Country	Method	Sensitivity	Specificity	Timeliness (weeks)
Greece	HMM	85%	100%	10
	PI	<50%	100%	25
	HMM-PI	91%	100%	6
Romania	HMM	61%	98%	0
	PI	67%	94%	35
	HMM-PI	100%	77%	0

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.
- PI performed better in both countries with:

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.
- PI performed better in both countries with:
    - 100% sensitivity,

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.
- PI performed better in both countries with:
  - 100% sensitivity,
  - no delay in triggering an alarm but

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.
- PI performed better in both countries with:
    - 100% sensitivity,
    - no delay in triggering an alarm but
    - moderate specificity ($< 70\%$).

Results: Post-epidemic state

- When applied for detecting post-epidemic state, no method reached a satisfying balance among all metrics.
- PI performed better in both countries with:
  - 100% sensitivity,
  - no delay in triggering an alarm but
  - moderate specificity ($< 70\%$).

- When applied retrospectively (i.e., at the end of each state), the simple two-state HMM gave almost perfect performance in terms of sensitivity ($\approx 100\%$) and specificity ($\approx 100\%$).

Conclusions

- No method is a panacea for detecting in real time all types of HIV outbreaks and/or epidemic states.

Conclusions

- No method is a panacea for detecting in real time all types of HIV outbreaks and/or epidemic states.
- We recommend HMM-PI and PI for monitoring large and abrupt outbreaks that do not return to pre-epidemic levels.

Conclusions

- No method is a panacea for detecting in real time all types of HIV outbreaks and/or epidemic states.
- We recommend HMM-PI and PI for monitoring large and abrupt outbreaks that do not return to pre-epidemic levels.
- Classic control charts are adequate for detecting the start and/or the end of very small epidemics.

Conclusions

- No method is a panacea for detecting in real time all types of HIV outbreaks and/or epidemic states.
- We recommend HMM-PI and PI for monitoring large and abrupt outbreaks that do not return to pre-epidemic levels.
- Classic control charts are adequate for detecting the start and/or the end of very small epidemics.
- Although HMM did not reach a high real-time performance, it may provide useful insights when applied retrospectively.

References

DES JARLAIS, D.C. et al. (2020). HIV Outbreaks Among People Who Inject Drugs in Europe, North America, and Israel. *Lancet HIV, 7(6), e434-e442*.

JOMBART, T. et al. (2021). Real-time monitoring of COVID-19 dynamics using automated trend fitting and anomaly detection. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376(1829), 20200266*.

LYTRAS, T. et al. (2019). FluHMM: A simple and flexible Bayesian algorithm for sentinel influenza surveillance and outbreak detection. *Statistical Methods in Medical Research*,28(6):1826-1840.

YUAN, M. et al. (2019). A systematic review of aberration detection algorithms used in public health surveillance. *Journal of Biomedical Informatics, 94, 103181*.

Acknowledgements

The present work was co-funded by the European Union and Greek national funds through the "Operational Program Competitiveness, Entrepreneurship and Innovation", under the call "RESEARCH-CREATE-INNOVATE" (project code: T1EΔK-03812).

With the co-financing of Greece and European Union

Thank You!